DOI: 10.1111/j.1541-0420.2010.01504.x

Nonparametric Comparison for Panel Count Data with Unequal Observation Processes

Xingqiu Zhao[1,*] **and Jianguo Sun**[2,**]

[1]Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China
[2]Department of Statistics, University of Missouri, 146 Middlebush Hall, Columbia, Missouri 65211, U.S.A.
[*]*email:* xingqiu.zhao@polyu.edu.hk
[**]*email:* sunj@missouri.edu

Summary. This article considers nonparametric comparison of several treatment groups based on panel count data, which often occur in, among others, medical follow-up studies and reliability experiments concerning recurrent events. For the problem, most of the existing procedures require that observation processes are identical across different treatment groups among other requirements. We propose a new class of nonparametric test procedures that allow different observation processes. The new test statistics are constructed based on the integrated weighted differences between the estimated mean functions of the underlying recurrent event processes. The asymptotic distributions of the proposed test statistics are established and their finite-sample properties are examined through Monte Carlo simulations, which indicate that the proposed approach works well for practical situations. An illustrative example is provided.

Key words: Counting processes; Medical follow-up study; Nonparametric comparison; Panel count data; Unequal observation scheme.

1. Introduction

Panel count data commonly refer to the data arising from studies concerning recurrent events in which each subject is observed only at several distinct time points instead of continuously (Kalbfleisch and Lawless, 1985). Examples of recurrent events include disease infections and machine failures (Cook, Lawless, and Nadeau, 1996; Cook and Lawless, 2007). For panel count data, no information is available on subjects between observation time points and only the numbers of the occurrences of the events between the observation times are known. Also the number of observations and the observation times may vary from subject to subject. The fields that often produce such data include clinical trials, medical follow-up studies, reliability experiments, sociological studies, and tumorgenicity experiments. In this article, we consider the nonparametric comparison of several treatment groups based on panel count data when the observation scheme or process defining the total number of observations and observation times may be different for different groups.

A well-known example of panel count data is given by a bladder cancer study conducted by the Veterans Administration Cooperative Urological Research Group (Sun and Wei, 2000; Wellner and Zhang, 2000). The study consists of three treatment groups: placebo, pyridoxine, and thiotepa, and a main objective was to compare the effects of the three treatments on the frequency of the bladder tumor recurrence. Several authors have noticed that the patients in the thiotepa group tended to visit the clinic centers more often than those in the placebo group. That is, the observation process seems to differ among the three treatment groups. Another more general situation that gives different observation processes is when observation times are hospitalization times of patients at which the occurrences of some recurrent events are determined.

Several nonparametric comparison procedures for panel count data have been developed in the literature. For example, one of the early procedures was given by Thall and Lachin (1988), who suggested transforming the problem to a multivariate comparison problem and then applying a multivariate Wilcoxon-type rank test. Sun and Fang (2003) proposed a nonparametric approach under the assumption that treatment indicators can be regarded as independent and identically distributed random variables and observation times have the same distribution for different groups. Also Zhang (2006) and Balakrishnan and Zhao (2009) proposed some nonparametric tests for the situation where the distributions of the total number of observations and observation times are identical for different treatment groups. The literature that discussed the analysis of panel count data also includes Hu, Lagakos, and Lockhart (2009), Sun and Kalbfleisch (1995), and Wellner and Zhang (2000), who studied nonparametric estimation of the mean function (MF) of the underlying recurrent event process. In addition, Hu, Sun, and Wei (2003), Sun and Wei (2000), Zhang (2002), Wellner and Zhang (2007), and Lu, Zhang, and Huang (2009) considered regression analysis of panel count data.

The remainder of the article is organized as follows. After introducing some notations and briefly reviewing the isotonic regression estimator for panel count data, Section 2 presents a class of nonparametric test statistics for comparing several treatment groups with respect to their MFs. The test procedures allow different observation processes for different treatment groups and the test statistics are formulated as the integrated weighted difference between the estimated MFs of the underlying recurrent event processes. The asymptotic normality of the presented test statistics is also established in Section 2. Section 3 presents some results obtained from a simulation study performed to assess the finite-sample properties of the proposed test procedure. In Section 4, we apply the proposed approach to the bladder tumor study discussed above and Section 5 contains some concluding remarks.

2. Treatment Comparison Based on Panel Count Data

Consider a recurrent event study that involves n independent subjects from k different treatment groups. Let n_l denote the number of subjects in the lth group and S_l the set of indices for subjects in group l, where $n_1 + \cdots + n_k = n$. Also let $N_i^{(l)}(t)$ denote the counting process representing the total number of occurrences of the recurrent event of interest up to time t from subject i in group l and $\Lambda_l(t) = E\{N_i^{(l)}(t)\}$, the common mean function of $N_i^{(l)}(t)$ for $i \in S_l, l = 1, \ldots, k$. Suppose that the goal is to test the hypothesis $H_0 : \Lambda_1(t) = \cdots = \Lambda_k(t)$. Also suppose that each subject is observed only at discrete time points and let $0 < T_{i,1}^{(l)} < \cdots < T_{i,K_i^{(l)}}^{(l)}$ denote the observation time points for subject i in group l with $K_i^{(l)}$ representing the total number of observation time points. That is, only panel count data are available. In the following text, we will assume that $\{(K_i^{(l)}; T_{i,1}^{(l)}, \ldots, T_{i,K_i^{(l)}}^{(l)}), i \in S_l\}$ are independent and identically distributed as $\{K^{(l)}; T_1^{(l)}, \ldots, T_{K^{(l)}}^{(l)}\}$ and are independent of the counting processes $N_i^{(l)}$'s. Also $\{N_i^{(l)}, i \in S_l\}$ are assumed to be identically distributed as $N^{(l)}$. Let $\mathbf{X}^{(l)} = (K^{(l)}, \mathbf{T}^{(l)}, \mathbf{N}^{(l)})$, where $K^{(l)}$ is a random variable representing the number of observations for a subject in group l in general, $\mathbf{T}^{(l)} = (T_1^{(l)}, \ldots, T_{K^{(l)}}^{(l)})$ and $\mathbf{N}^{(l)} = (N^{(l)}(T_1^{(l)}), \ldots, N^{(l)}(T_{K^{(l)}}^{(l)}))$. Then $\{\mathbf{X}_i^{(l)} = (K_i^{(l)}, \mathbf{T}_i^{(l)}, \mathbf{N}_i^{(l)}), i \in S_l\}$ is a random sample of size n_l from the distribution of $\mathbf{X}^{(l)}, l = 1, \ldots, k$.

Before constructing the test statistics for H_0, we first introduce the isotonic regression estimators of the MF Λ_l's (Sun and Kalbfleisch, 1995; Wellner and Zhang, 2000). Let $s_1^{(l)}, \ldots, s_{m_l}^{(l)}$ denote the ordered distinct observation times in the set $\{T_{i,j}^{(l)};\ j = 1, \ldots, K_i^{(l)},\ i \in S_l\}$ and $w_\ell^{(l)}$ and $\bar{N}_\ell^{(l)}$ be the number and mean value, respectively, of the observations made at time $s_\ell^{(l)}$ from subjects in group $l, \ell = 1, \ldots, m_l$. Then the isotonic regression estimator $\hat{\Lambda}_n^{(l)}(t)$ of $\Lambda_l(t)$ is defined as a nondecreasing step function with possible jumps at the $s_\ell^{(l)}$'s and has the form

$$\hat{\Lambda}_n^{(l)}\left(s_\ell^{(l)}\right) = \max_{r \leqslant \ell} \min_{s \geqslant \ell} \frac{\sum_{v=r}^{s} w_v^{(l)} \bar{N}_v^{(l)}}{\sum_{v=r}^{s} w_v^{(l)}} = \min_{s \geqslant \ell} \max_{r \leqslant \ell} \frac{\sum_{v=r}^{s} w_v^{(l)} \bar{N}_v^{(l)}}{\sum_{v=r}^{s} w_v^{(l)}}, \quad \ell = 1, \ldots, m_l, \tag{1}$$

which is the isotonic regression of the $\bar{N}_\ell^{(l)}$'s with weights $w_\ell^{(l)}$'s (Robertson, Wright, and Dykstra, 1988).

To present the test statistics for the hypothesis H_0, let τ denote the largest observation time, $p_l = n_l/n, \pi_l$ be the limit of p_l. Define

$$G_l(t) = \mathrm{E}\left[\sum_{j=1}^{K^{(l)}} I\left\{T_j^{(l)} \leqslant t\right\}\right], \qquad g_l(t) = G_l'(t),$$

$$G(t) = \sum_{l=1}^{k} \pi_l G_l(t), \quad g(t) = G'(t) = \sum_{l=1}^{k} \pi_l g_l(t),$$

$$\nu_l(t) = g(t)/g_l(t),$$

$$G_n^{(l)}(t) = \frac{1}{n_l} \sum_{i \in S_l} \sum_{j=1}^{K_i^{(l)}} I\left(T_{i,j}^{(l)} \leqslant t\right),$$

the empirical observation process from group l, and

$$G_n(t) = \sum_{l=1}^{k} p_l G_n^{(l)}(t),$$

the overall empirical observation process. Also define

$$\Psi_n^{(l)} = \int_0^\tau W_n(t) \hat{\Lambda}_n^{(l)}(t)\, dG_n(t), \tag{2}$$

a summary measure of the event history in group l, and

$$\hat{\sigma}_l^2 = \frac{1}{n_l} \sum_{i \in S_l} \left[\sum_{j=1}^{K_i^{(l)}} A_n^{(l)}\left(T_{i,j}^{(l)}\right)\left\{N_i^{(l)}\left(T_{i,j}^{(l)}\right) - \hat{\Lambda}_n^{(l)}\left(T_{i,j}^{(l)}\right)\right\}\right]^2,$$

where the $W_n(t)$'s are bounded weight processes and

$$A_n^{(l)}(t) = \sum_{r=1}^{k} \frac{n_r}{n} W_n(t) \frac{G_n^{(r)}(t) - G_n^{(r)}(t-)}{G_n^{(l)}(t) - G_n^{(l)}(t-)},$$

$l = 1, \ldots, k$. Then we propose to use the following test statistic

$$T = \sum_{l=1}^{k} c_l \left(\Psi_n^{(l)} - \bar{\Psi}_n\right)^2,$$

where $c_l = n_l/\hat{\sigma}_l^2$ and $\bar{\Psi}_n = \sum_{l=1}^k \alpha_l \Psi_n^{(l)}$ with $\alpha_l = c_l/C$ and $C = \sum_{l=1}^k c_l$.

It is easy to see that the test statistic T represents the integrated weighted difference among the estimated MF $\hat{\Lambda}_n^{(l)}$ and the resulting test is analogous to Welch's test in analysis of variance with unequal variances. The similar statistics are also commonly used in failure time data analysis (Kalbfleisch and Prentice, 2002) and panel count data analysis (Sun and Fang, 2003; Zhang, 2006). In particular, when observation processes are identical across different treatment groups, Zhang (2006) proposed to use the test statistic

$$T_{1Z}^{(l)} = \sqrt{\frac{n_1 n_l}{n}}\left\{\Psi_n^{(1)} - \Psi_n^{(l)}\right\}$$

to test the hypothesis H_0. For the situation where the observation process may differ across different groups, in the discussion section, Zhang (2006) suggested the test statistic

$$T_{2Z}^{(l)} = \sqrt{\frac{n_1 n_l}{n}} \left\{ \int \hat{\Lambda}_n^{(1)}(t) \frac{1}{\hat{g}_1(t)} dG_n^{(1)}(t) - \int \hat{\Lambda}_n^{(l)}(t) \frac{1}{\hat{g}_l(t)} dG_n^{(l)}(t) \right\},$$

where $\hat{g}_l(t)$ represents a kernel-smoothed estimator of $g_l(t)$. It should be noted that $T_{1Z}^{(l)}$ and $T_{2Z}^{(l)}$ do not give the same test procedure when the observation processes are identical and the latter involves estimation of some derivative functions. More importantly, the properties of $T_{2Z}^{(l)}$ are still unknown.

To derive the asymptotic distribution of the test statistic T, first we need to define some new notations and establish the asymptotic normality of the functional of the isotonic regression estimator. Let $\Lambda_0(t)$ denote the common MFs of the $N_i^{(l)}(t)$'s under H_0 and Λ_0^{-1} the inverse function of Λ_0. Also let $A \circ \Lambda_0^{-1}$ denote the composition of two functions A and Λ_0^{-1},

$$\Psi_n^{(0)} = \int_0^\tau W_n(t) \Lambda_0(t)\, dG_n(t),$$

and $\Delta_n^{(l)} = \sqrt{n_l}(\Psi_n^{(l)} - \Psi_n^{(0)})$. The following theorem gives the asymptotic normality of the functional of the isotonic regression estimator.

Theorem 1: *Suppose that $A_l(t)$ is a bounded weight process with $A_l \circ \Lambda_0^{-1}$ being a bounded Lipschitz function and define*

$$V_n^{(l)} = \sqrt{n_l} \int_0^\tau A_l(t) \left\{ \hat{\Lambda}_n^{(l)}(t) - \Lambda_0(t) \right\} dG_l(t)$$

and

$$\bar{V}_n^{(l)} = \frac{1}{n_l} \sum_{i \in S_l} \sum_{j=1}^{K_i^{(l)}} A_l\left(T_{i,j}^{(l)}\right) \left\{ N_i^{(l)}\left(T_{i,j}^{(l)}\right) - \Lambda_0\left(T_{i,j}^{(l)}\right) \right\}.$$

Then under H_0 and the Conditions 1–3 given in the Appendix and as $n \to \infty$, we have that $V_n^{(l)} = \bar{V}_n^{(l)} + o_p(1)$ and both $V_n^{(l)}$ and $\bar{V}_n^{(l)}$ converge in distribution to the normal random variable V_l with mean zero and variance

$$\sigma_l^2 = \mathrm{E}\left[\left(\sum_{j=1}^{K_i^{(l)}} A_l\left(T_{i,j}^{(l)}\right) \left\{ N_i^{(l)}\left(T_{i,j}^{(l)}\right) - \Lambda_0\left(T_{i,j}^{(l)}\right) \right\} \right)^2 \right]. \quad (3)$$

The theorem given above can be proved by using the same techniques as those used in the proof of Theorem 2.1 of Balakrishnan and Zhao (2009) and modifying the proof of Theorem 2 of Zhang (2006). Thus the proof is omitted. It is worth pointing out that here we do not need the monotone assumption for the function $A_l \circ \Lambda_0^{-1}$ required by Theorem 2 of Zhang (2006). In the following theorem, we establish the asymptotic distribution of $\Delta_n^{(l)}$.

Theorem 2: *Assume that $W(t)$ is a bounded function such that*

$$\int_0^\tau \{W_n(t) - W(t)\}^2 \, dG_l(t) = o_p(n^{-1/3}) \quad (4)$$

and $A_l \circ \Lambda_0^{-1}$ is a bounded Lipschitz function with $A_l(t) = W(t)\nu_l(t), l = 1, \ldots, k$. Also assume that the Conditions 1–3 given in the Appendix hold and g_l is bounded with a positive lower bound on $[\tau_0, \tau]$. Then as $n \to \infty$ and under H_0, we have

$$\Delta_n^{(l)} = \bar{V}_n^{(l)} + o_p(1)$$

with the $\bar{V}_n^{(l)}$'s defined in Theorem 1.

It follows from Theorems 1 and 2 that $\Delta_n^{(l)}$ asymptotically follows a normal distribution. We remark that if observation processes are identical across different groups, then we have $G_l = G$ and $g_l = g$ for all l. In this case, the result described above is the same as that given in Theorem 3 of Zhang (2006) but without the monotone assumption for the weight process. The next theorem shows that one can consistently estimate σ_l^2 by $\hat{\sigma}_l^2$ given above.

Theorem 3: *Suppose that all conditions described in Theorem 2 and the Condition A4 given in the Appendix hold. Also suppose that the g_l''s are bounded and*

$$\max_{i \in S_l} \mathrm{E}\left[\sum_{j=1}^{K_i^{(l)}} \left\{ W_n\left(T_{i,j}^{(l)}\right) - W\left(T_{i,j}^{(l)}\right) \right\}^2 \right] \to 0. \quad (5)$$

Then as $n \to \infty$, we have that $\hat{\sigma}_l^2 \to_p \sigma_l^2$.

The proofs of Theorems 2 and 3 are sketched in the Appendix. These two theorems suggest that the testing of the hypothesis H_0 can be carried out by applying the statistic T based on the central χ^2-distribution with $(k-1)$ degrees of freedom. Note that this test procedure is valid no matter whether observation processes are identical or not among different treatment groups.

To employ the test procedure proposed above, one needs to choose the weight process $W_n(t)$. For this, a simple and natural choice is clearly $W_n^{(1)}(t) = 1$. Another natural choice is

$$W_n^{(2)}(t) = \frac{1}{n} \sum_{l=1}^{k} \sum_{i \in S_l} I\left(t \leqslant T_{i,K_i^{(l)}}^{(l)} \right).$$

It is easy to see that the first weight function weights everything equally, whereas the second one assigns the weight according to the number of subjects under observation. Of course, one could also use $W_n^{(3)}(t) = 1 - W_n^{(2)}(t)$. It is easy to verify that all three weight processes satisfy the conditions required in Theorem 3.

3. A Simulation Study

An extensive simulation study was conducted to assess the finite-sample properties of the test procedure proposed in the previous section. In the study, we focused on the two-sample comparison problem with $k = 2$ and in this case, the test statistic has the form

$$T = \frac{n_1}{\hat{\sigma}_1^2}\left(\Psi_n^{(1)} - \bar{\Psi}_n\right)^2 + \frac{n_2}{\hat{\sigma}_2^2}\left(\Psi_n^{(2)} - \bar{\Psi}_n\right)^2$$

with

$$\bar{\Psi}_n = \left\{ \frac{n_1}{\hat{\sigma}_1^2} \Psi_n^{(1)} + \frac{n_2}{\hat{\sigma}_2^2} \Psi_n^{(2)} \right\} \Big/ \left\{ \frac{n_1}{\hat{\sigma}_1^2} + \frac{n_2}{\hat{\sigma}_2^2} \right\}.$$

For the generation of panel count data $\{K_i^{(l)}, T_{i,j}^{(l)}, N_i^{(l)}(T_{i,j}^{(l)}), j = 1, \ldots, K_i^{(l)}, i \in S_l, l = 1, 2\}$, we first generated the number of observation times $K_i^{(l)}$ based on the uniform distribution over $\{1, \ldots, b_l\}$, where b_1 and b_2 are some integers. That is, $\Pr(K_i^{(l)} = x) = 1/b_l, x = 1, \ldots, b_l$. Given $K_i^{(l)}$, the observation times $T_{i,j}^{(l)}$'s were generated as the order statistics of $K_i^{(l)}$ random variables from the probability density function

$$f(x;\theta_l) = \frac{\theta_l + 1}{\tau^{\theta_l+1} - 1} x^{\theta_l}, \quad 1 \leqslant x \leqslant \tau,$$

where θ_1 and θ_2 are some constants. The constants b_l's and θ_l's are used to control the observation processes.

For given $K_i^{(l)}$'s and $T_{i,j}^{(l)}$'s, the panel counts $N_i^{(l)}(T_{i,j}^{(l)})$'s were generated from the Poisson processes with one of the following two different types of conditional MFs

MF 1. $\Lambda_i^{(1)}(t \,|\, \nu_i^{(1)}) = \nu_i^{(1)} t$ for $i \in S_1, \Lambda_i^{(2)}(t \,|\, \nu_i^{(2)}) = \nu_i^{(2)} t \times \exp(\beta)$ for $i \in S_2$,

MF 2. $\Lambda_i^{(1)}(t \,|\, \nu_i^{(1)}) = \nu_i^{(1)} t$ for $i \in S_1, \Lambda_i^{(2)}(t \,|\, \nu_i^{(2)}) = \nu_i^{(2)} \sqrt{\beta t}$ for $i \in S_2$.

Here $\{\nu_i^{(l)}, i \in S_l, l = 1, 2\}$ are latent variables that were set to be equal to one or generated from the Gamma$(2, 1/2)$ distribution. The former means that the $N_i^{(l)}(t)$'s are Poisson processes, whereas the latter gives mixed Poisson processes. Note that we can write $N_i^{(l)}(T_{i,j}^{(l)})$ as

$$\begin{aligned} N_i^{(l)}\big(T_{i,j}^{(l)}\big) &= N_i^{(l)}\big(T_{i,1}^{(l)}\big) + \big\{N_i^{(l)}\big(T_{i,2}^{(l)}\big) - N_i^{(l)}\big(T_{i,1}^{(l)}\big)\big\} \\ &\quad + \cdots + \big\{N_i^{(l)}\big(T_{i,j}^{(l)}\big) - N_i^{(l)}\big(T_{i,j-1}^{(l)}\big)\big\} \\ &\equiv \xi_{i,1}^{(l)} + \cdots + \xi_{i,j}^{(l)} \end{aligned}$$

with $\{\xi_{i,j}^{(l)}, j = 1, \ldots, K_i^{(l)}\}$ being independent Poisson random variables with means

$$\big\{\Lambda_i^{(l)}\big(T_{i,j}^{(l)} \,\big|\, \nu_i^{(l)}\big) - \Lambda_i^{(l)}\big(T_{i,j-1}^{(l)} \,\big|\, \nu_i^{(l)}\big), j = 1, \ldots, K_i^{(l)}\big\}.$$

Thus for given $K_i^{(l)}$'s, $T_{i,j}^{(l)}$'s, and $\nu_i^{(l)}$'s, the panel counts $N_i^{(l)}(T_{i,j}^{(l)})$'s can be generated by generating the $\xi_{i,j}^{(l)}$'s.

To give an idea about the shapes of the conditional MFs defined above, Figures 1 and 2 display them for the cases of $\nu_i^{(l)} = 1$ with three different values of β. It can be seen that the functions in MF 1 do not overlap with each other, whereas those in MF 2 cross over each other. The results given below are based on 1000 replications and for the case of $n_1 = n_2 = 50$ or 100 except in Table 5.

Tables 1 and 2 give the estimated sizes and powers of the proposed test procedures at the significance level $\alpha = 0.05$ with different values of $\beta, (b_1, b_2)$ and (θ_1, θ_2) and the use of MF 1. The results in Table 1 are for situations where the underlying recurrent event processes were Poisson processes and Table 2 corresponds to situations where the panel count data were generated from mixed Poisson processes. Here we considered all three weight processes discussed in the previous section and took $\tau = 10$. The results indicate that the proposed test procedure seems to have proper size and good power to detect treatment differences. In particular, the procedure with the weight process $W_n^{(1)}(t)$ seems to uniformly outperform the procedures with other two weight processes although the power is quite close among them. Also as expected, the power increased when the sample size increased and the power decreased in the presence of more variability.

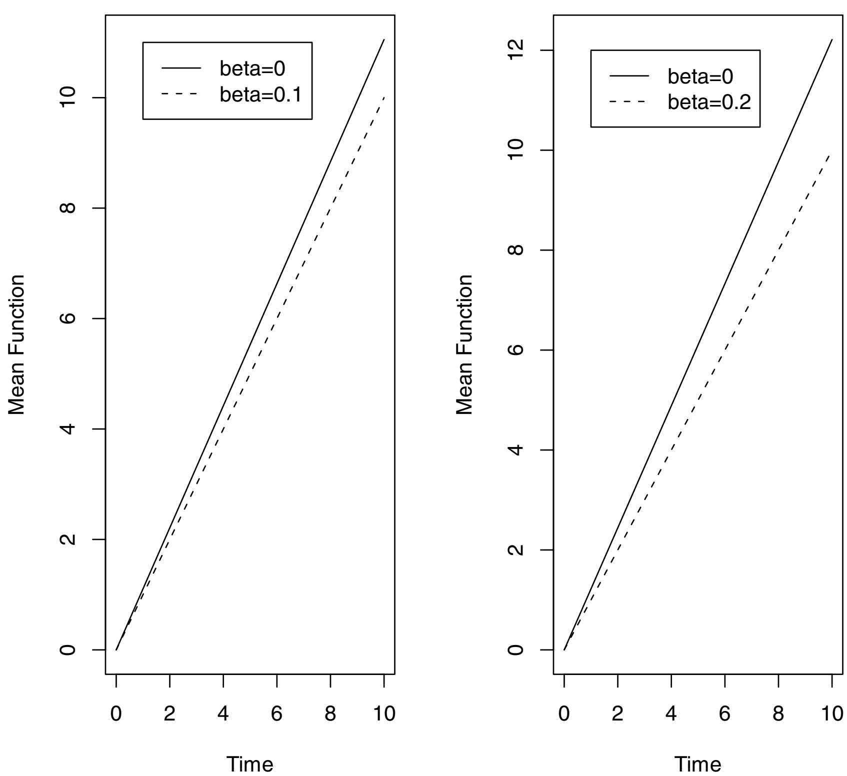

Figure 1. MF 1 with $\nu_i^{(l)} = 1$ and $\beta = 0.1, 0.2$.

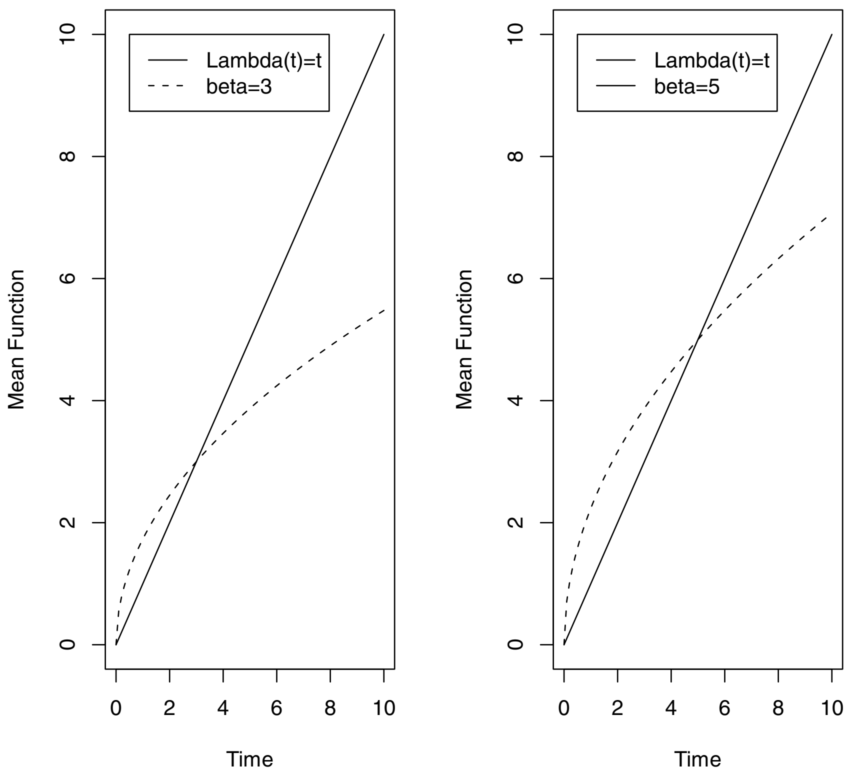

Figure 2. MF 2 with $\nu_i^{(l)} = 1$ and $\beta = 3, 5$.

Table 1

Estimated size and power of the proposed test procedure based on the panel count data generated from Poisson processes with MF 1

			$n_1 = n_2 = 50$			$n_1 = n_2 = 100$		
(b_1, b_2)	(θ_1, θ_2)	β	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$
(6,8)	(0.0,0.1)	0.0	0.056	0.051	0.051	0.052	0.049	0.048
		0.1	0.256	0.230	0.246	0.458	0.423	0.410
		0.2	0.759	0.710	0.719	0.949	0.931	0.941
		0.3	0.981	0.961	0.971	1.000	1.000	1.000
	(0.0,0.5)	0.0	0.056	0.053	0.049	0.049	0.058	0.044
		0.1	0.293	0.273	0.266	0.450	0.425	0.431
		0.2	0.772	0.732	0.734	0.957	0.934	0.953
		0.3	0.983	0.967	0.973	1.000	1.000	1.000
(8,6)	(0.0,0.1)	0.0	0.050	0.046	0.059	0.051	0.049	0.049
		0.1	0.275	0.262	0.251	0.455	0.422	0.441
		0.2	0.750	0.716	0.733	0.961	0.938	0.934
		0.3	0.980	0.965	0.979	1.000	1.000	1.000
	(0.0,0.5)	0.0	0.052	0.056	0.055	0.055	0.051	0.056
		0.1	0.271	0.249	0.256	0.456	0.438	0.452
		0.2	0.737	0.700	0.726	0.957	0.941	0.942
		0.3	0.975	0.958	0.968	1.000	1.000	1.000

The estimated powers of the proposed test procedures obtained with the use of MF 2 are presented in Tables 3 and 4 with all other setups being the same as in Tables 1 and 2, respectively. Note that for this situation, only the estimated powers are given and it can be easily seen that as expected, the power was lower than that for the situations considered in Tables 1 and 2. In particular, the results suggest that the power of the proposed test procedure depends on the weight process when the underlying MFs cross over each other. In other words, some knowledge about the shapes of the underlying MFs can be used or would be needed to select a better weight process.

Table 2
Estimated size and power of the proposed test procedure based on the panel count data generated from mixed Poisson processes with MF 1

			$n_1 = n_2 = 50$			$n_1 = n_2 = 100$		
(b_1, b_2)	(θ_1, θ_2)	β	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$
(6,8)	(0.0,0.1)	0.0	0.054	0.050	0.058	0.053	0.052	0.053
		0.1	0.089	0.092	0.095	0.130	0.133	0.113
		0.2	0.209	0.196	0.188	0.335	0.325	0.305
		0.3	0.387	0.387	0.339	0.642	0.633	0.554
	(0.0,0.5)	0.0	0.057	0.054	0.049	0.046	0.048	0.051
		0.1	0.099	0.095	0.099	0.112	0.104	0.103
		0.2	0.217	0.208	0.189	0.352	0.353	0.307
		0.3	0.429	0.427	0.383	0.654	0.652	0.568
(8,6)	(0.0,0.1)	0.0	0.050	0.052	0.056	0.055	0.052	0.055
		0.1	0.113	0.100	0.111	0.145	0.137	0.130
		0.2	0.208	0.189	0.181	0.348	0.330	0.295
		0.3	0.436	0.433	0.384	0.642	0.632	0.569
	(0.0,0.5)	0.0	0.052	0.050	0.053	0.049	0.048	0.050
		0.1	0.089	0.090	0.088	0.141	0.130	0.124
		0.2	0.219	0.206	0.198	0.353	0.349	0.309
		0.3	0.402	0.400	0.367	0.677	0.656	0.596

Table 3
Estimated power of the proposed test procedure based on the panel count data generated from Poisson processes with MF 2

			$n_1 = n_2 = 50$			$n_1 = n_2 = 100$		
(b_1, b_2)	(θ_1, θ_2)	β	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$
(6,8)	(0.0,0.1)	3	0.976	0.875	1.000	1.000	0.984	1.000
		5	0.216	0.087	0.815	0.335	0.100	0.972
		8	0.491	0.711	0.079	0.743	0.927	0.075
	(0.0,0.5)	3	0.984	0.910	1.000	1.000	1.000	1.000
		5	0.256	0.104	0.832	0.428	0.142	0.978
		8	0.368	0.609	0.078	0.614	0.882	0.072
(8,6)	(0.0,0.1)	3	0.963	0.835	1.000	1.000	0.986	1.000
		5	0.202	0.084	0.804	0.345	0.094	0.984
		8	0.467	0.683	0.075	0.710	0.939	0.069
	(0.0,0.5)	3	0.977	0.908	1.000	1.000	0.995	1.000
		5	0.284	0.124	0.842	0.412	0.126	0.983
		8	0.391	0.626	0.085	0.664	0.897	0.070

As pointed out by a referee, a question of practical interest is how the proposed test performs compared to that given in Zhang (2006) for the situation where observation processes are the same. Another natural question that one may ask is if one could simply apply the test procedures developed under the identical observation process assumption to unequal observation process situations. To answer these, we also generated panel count data with equal and unequal observation processes, respectively, and compared the sizes and powers of Zhang's test procedure (2006) and the test procedure proposed here. Table 5 presents the estimated sizes and powers of the two test procedures under the same setups as those in Tables 1 and 3 and with $n_1 = 30$ and $n_2 = 40$. Here, suggested by the associate editor, the sample sizes were chosen based on the example discussed in the next section. The results indicate that the two methods gave similar performance when the observation processes are the same. However, when the observation processes differ between treatment groups, Zhang's procedure, which does not take into account the difference, could overestimate the size.

In addition to assessing the size and power of the proposed test procedure, we also studied the chi-square distribution approximation to the distribution of the proposed test statistics by employing quantile plots of the test statistic T against the $\chi^2(1)$ distribution. The results suggest that the chi-square approximation is quite good under the situations considered in Tables 1–4. In summary, the simulation results suggest that the proposed test procedure seems to work well for practical situations. With respect to the selection of weight processes, $W_n^{(1)}$ seems to be the best choice if the MFs of the underlying counting processes do not cross over each other. On the other hand, when the MFs cross over each other, some knowledge

Table 4

Estimated power of the proposed test procedure based on the panel count data generated from mixed Poisson processes with MF 2

			$n_1 = n_2 = 50$			$n_1 = n_2 = 100$		
(b_1, b_2)	(θ_1, θ_2)	β	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$
(6,8)	(0.0,0.1)	3	0.487	0.349	0.698	0.727	0.557	0.937
		5	0.097	0.066	0.252	0.123	0.068	0.408
		8	0.147	0.221	0.069	0.228	0.396	0.061
	(0.0,0.5)	3	0.477	0.380	0.687	0.778	0.633	0.925
		5	0.095	0.070	0.240	0.133	0.087	0.424
		8	0.109	0.167	0.062	0.171	0.300	0.058
(8,6)	(0.0,0.1)	3	0.475	0.359	0.709	0.728	0.573	0.919
		5	0.088	0.062	0.241	0.120	0.065	0.419
		8	0.155	0.221	0.064	0.202	0.353	0.064
	(0.0,0.5)	3	0.527	0.416	0.708	0.780	0.610	0.942
		5	0.095	0.060	0.275	0.127	0.068	0.434
		8	0.118	0.172	0.065	0.182	0.306	0.053

Table 5

Estimated sizes and powers of the proposed test procedure and Zhang (2006)*'s procedure with* $n_1 = 30, n_2 = 40$

			Proposed procedure			Zhang (2006)'s procedure		
(b_1, b_2)	(θ_1, θ_2)	β	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$	$W_n^{(1)}$	$W_n^{(2)}$	$W_n^{(3)}$
					Equal observation processes			
(6,6)	(0.0,0.0)	0.0	0.053	0.053	0.058	0.052	0.052	0.057
		0.1	0.190	0.179	0.198	0.180	0.171	0.169
		0.2	0.568	0.530	0.549	0.563	0.520	0.536
		0.3	0.895	0.867	0.895	0.895	0.866	0.896
		3	0.855	0.640	0.982	0.867	0.657	0.990
		5	0.140	0.066	0.608	0.143	0.062	0.630
		8	0.361	0.595	0.067	0.333	0.574	0.062
					Unequal observation processes			
(6,8)	(0.0,0.0)	0.0	0.047	0.048	0.052	0.069	0.072	0.075
	(0.0,0.1)		0.048	0.051	0.050	0.076	0.071	0.089
	(0.0,0.5)		0.059	0.055	0.058	0.081	0.081	0.082
	(0.0,0.1)	0.1	0.187	0.177	0.182	0.180	0.169	0.170
		0.2	0.560	0.542	0.551	0.555	0.518	0.540
		0.3	0.898	0.864	0.874	0.910	0.875	0.877
		3	0.860	0.648	0.986	0.875	0.687	0.993
		5	0.135	0.062	0.609	0.154	0.066	0.648
		8	0.331	0.502	0.061	0.304	0.473	0.060
	(0.0,0.5)	0.1	0.193	0.186	0.183	0.197	0.182	0.178
		0.2	0.584	0.547	0.518	0.578	0.548	0.525
		0.3	0.912	0.890	0.873	0.904	0.890	0.881
		3	0.902	0.769	0.992	0.914	0.802	0.997
		5	0.173	0.082	0.616	0.193	0.090	0.701
		8	0.273	0.445	0.050	0.243	0.416	0.057

about their shapes is needed for selecting a weight process that gives the better power. In this case, if the main difference of the MF occurs at earlier time periods, the weight process $W_n^{(2)}$ tends to be a better choice. However, one may want to employ $W_n^{(3)}$ if the main difference lays over later time periods.

4. An Application

In this section, we apply the proposed methodology to the bladder tumor study discussed before. The study consists of 116 bladder cancer patients and three treatments: placebo (47), pyridoxine (31), and thiotepa (38). All patients in the study had superficial bladder tumors when they entered the study and these tumors were removed before the start of the treatments. At each follow-up visit to the clinic office, the patient was examined and the number of recurrences of the bladder tumors since the previous visit was recorded with the tumors removed. Thus we have a set of panel count data $\{K_i^{(l)}, T_{i,j}^{(l)}, N_i^{(l)}(T_{i,j}^{(l)})\}$ with $k = 3$, where $K_i^{(l)}, T_{i,j}^{(l)}$ and $N_i^{(l)}(T_{i,j}^{(l)})$ denote, respectively, the total number of clinic

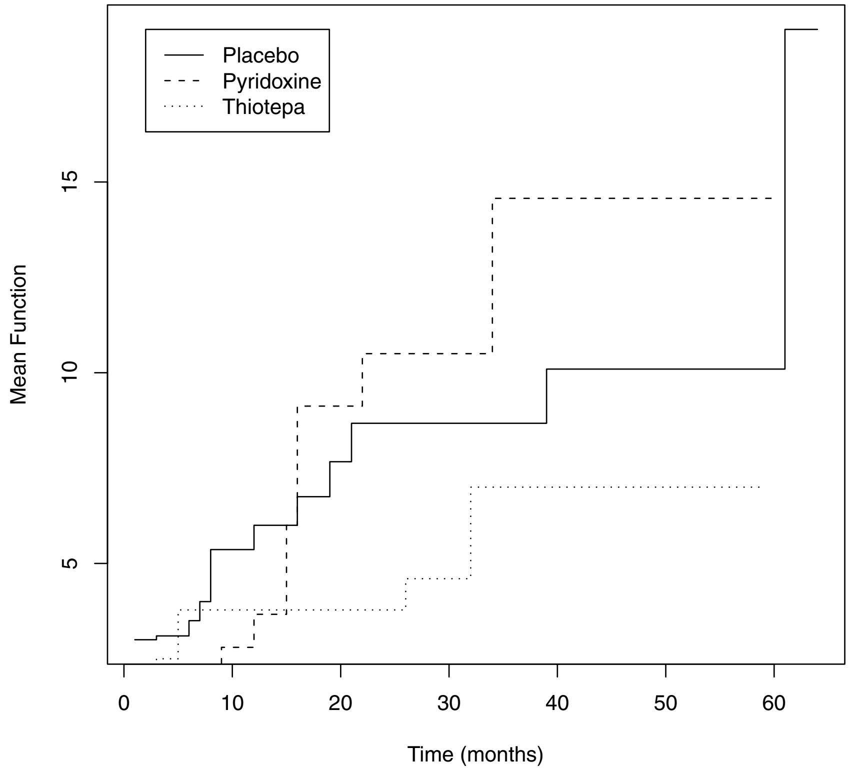

Figure 3. Estimates of the MFs for the bladder tumor study.

visits, the visit time, and the total number of recurrences of the bladder tumors up to time $T_{i,j}^{(l)}$ for patient i in treatment group l.

To compare the three treatment groups with respect to the recurrence rates of bladder tumors, define $\Lambda_1(t), \Lambda_2(t)$, and $\Lambda_3(t)$ to be the MFs corresponding to the placebo, pyridoxine, and thiotepa treatments, respectively. The application of the test procedure proposed in the previous sections yielded $T = 5.2805$ and the p-value of 0.0713 for testing H_0 with the use of weight process $W_n^{(1)}$. If using the weight process $W_n^{(2)}$ or $W_n^{(3)}$, we obtained $T = 0.0379$ or 21.7701 along with the p-value of 0.9812 or 0.00002, respectively. To understand these results, we calculated and plotted in Figure 3 the isotonic regression estimators of the three MFs. It can be seen from the figure that there exists some crossing during the initial period and this suggests that we should rely on the weight processes $W_n^{(1)}$ and $W_n^{(3)}$ but not $W_n^{(2)}$. The results given by $W_n^{(1)}$ and $W_n^{(3)}$ indicate that the recurrence rates of the bladder tumors were significantly different among the three treatment groups. In comparison, the test procedure proposed in Zhang (2006) gave the p-values of 0.0851, 0.1445, and 0.0840, respectively, based on the same three weight processes.

5. Concluding Remarks

This article discussed the problem of the multisample comparison of point processes when only panel count data are available and observation processes may be different across different samples. A class of nonparametric test procedures was proposed for the problem and the asymptotic properties of the test statistics were established. An extensive simulation study was carried out and suggested that the proposed approach works well for practical situations. As shown in both the simulation study and the application, in the presence of different observation processes, one needs to be careful in choosing a nonparametric comparison procedure as many existing procedures apply only to situations where the observation processes are identical across different treatment groups (Thall and Lachin, 1988; Sun and Fang, 2003; Zhang, 2006; Balakrishnan and Zhao, 2009).

To construct the proposed test statistics, we employed the isotonic regression estimators of the underlying MFs. As an alternative, one could instead use the maximum likelihood estimators. Wellner and Zhang (2000) showed that the nonparametric maximum likelihood estimator of the MFs could be more efficient than the isotonic regression or nonparametric maximum pseudo-likelihood estimator. Balakrishnan and Zhao (2009) also showed that the tests based on the nonparametric maximum likelihood estimator could be more powerful than these based on the nonparametric maximum pseudo-likelihood estimator for panel count data with identical observation processes. Thus as a future research direction, one may develop test procedures similar to the one developed here but by using the nonparametric maximum likelihood estimator.

Acknowledgements

We thank the editor, Dr David Zucker, the associate editor, and two referees for their many helpful and thoughtful comments and suggestions that greatly improved this article. The research reported here was supported in part by a grant from the Research Grant Council of Hong Kong.

References

Balakrishnan, N. and Zhao, X. (2009). New multi-sample nonparametric tests for panel count data. *The Annals of Statistics* **37,** 1112–1149.

Cook, R. J. and Lawless, J. F. (2007). *The Analysis of Recurrent Event Data*. New York: Springer-Verlag.

Cook, R. J., Lawless, J. F., and Nadeau, C. (1996). Robust tests for treatment comparisons based on recurrent event response. *Biometrics* **52,** 557–571.

Hu, X. J., Sun, J., and Wei, L. J. (2003). Regression parameter estimation from panel counts. *Scandinavian Journal of Statistics* **30,** 25–43.

Hu, X. J., Lagakos, S. W., and Lockhart, R. A. (2009). Generalized least squares estimation of the mean function of a counting process based on panel counts. *Statistica Sinica* **19,** 561–580.

Kalbfleisch, J. D. and Lawless, J. F. (1985). The analysis of panel data under a Markov assumption. *Journal of the American Statistical Association* **80,** 863–871.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. Hoboken, New Jersey: John Wiley & Sons.

Lu, M., Zhang, Y., and Huang, J. (2009). Semiparametric estimation methods for panel count data using monotone B-splines. *Journal of the American Statistical Association* **104,** 1060–1070.

Robertson, T., Wright, F. T., and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*. New York: John Wiley and Sons.

Sun, J. and Fang, H. B. (2003). A nonparametric test for panel count data. *Biometrika* **90,** 199–208.

Sun, J. and Kalbfleisch, J. D. (1995). Estimation of the mean function of point processes based on panel data. *Statistica Sinica* **5,** 279–289.

Sun, J. and Wei, L. J. (2000). Regression analysis of panel count data with covariate-dependent observation and censoring times. *Journal of the Royal Statistical Society, Series B* **62,** 293–302.

Thall, P. F. and Lachin, J. M. (1988). Analysis of recurrent events: Nonparametric methods for random-interval count data. *Journal of the American Statistical Association* **83,** 339–347.

Wellner, J. A. and Zhang, Y. (2000). Two estimators of the mean of a counting process with panel count data. *The Annals of Statistics* **28,** 779–814.

Wellner, J. A. and Zhang, Y. (2007). Two likelihood-based semiparametric estimation methods for panel count data with covariates. *The Annals of Statistics* **35,** 2106–2142.

Zhang, Y. (2002). A semiparametric pseudolikelihood estimation method for panel count data. *Biometrika* **89,** 39–48.

Zhang, Y. (2006). Nonparametric *k*-sample tests with panel count data. *Biometrika* **93,** 777–790.

Received December 2009. *Revised June* 2010.
Accepted August 2010.

Appendix

Proofs of Theorems 2 and 3

In this Appendix, we will use the same notation defined in the previous sections. Before proving Theorems 2 and 3, we need to describe four conditions.

Condition A1: *The MF Λ_0 is strictly increasing such that $\Lambda_0(\tau) \leqslant M$ for some constant $M \in (0, \infty)$;*

Condition A2: *There exists a constant K_0 such that $\Pr\{K^{(l)} \leqslant K_0\} = 1$ and that the random variables $T_{i,j}^{(l)}$'s take values in a bounded set $[\tau_0, \tau]$, where $0 < \tau_0 < \tau < \infty$;*

Condition A3: $\Pr\{\limsup_{n\to\infty} \max_i N_i^{(l)}(\tau) < \infty\} = 1$ *and $\mathrm{E}((N^{(l)}(t))^4) \leqslant M_1$ for all $t \leqslant \tau$, where M_1 is a constant.*

Condition A4: $\Pr\{\min_{1\leqslant l\leqslant k} \min_{1\leqslant j\leqslant K^{(l)}} (T_j^{(l)} - T_{j-1}^{(l)}) \geqslant s_0\} = 1$ *for some fixed time point s_0, where $T_0^{(l)} = 0$ and s_0 can be considered as the smallest length of consecutive observation times.*

Proof of Theorem 2. To prove this theorem, first note that we can rewrite it as

$$\begin{aligned}
\Delta_n^{(l)} &= \sqrt{n_l} \int W_n(t) \left\{\hat{\Lambda}_n^{(l)}(t) - \Lambda_0(t)\right\} \sum_{r=1}^{k} p_r \, dG_n^{(r)}(t) \\
&= \sum_{r=1}^{k} p_r \sqrt{n_l} \int \{W_n(t) - W(t)\} \left\{\hat{\Lambda}_n^{(l)}(t) - \Lambda_0(t)\right\} d \\
&\quad \times \left\{G_n^r(t) - G_r(t)\right\} \\
&\quad + \sum_{r=1}^{k} p_r \sqrt{n_l} \int \{W_n(t) - W(t)\} \left\{\hat{\Lambda}_n^{(l)}(t) - \Lambda_0(t)\right\} dG_r(t) \\
&\quad + \sum_{r=1}^{k} p_r \sqrt{n_l} \int W(t) \left\{\hat{\Lambda}_n^{(l)}(t) - \Lambda_0(t)\right\} d \\
&\quad \times \left\{G_n^{(r)}(t) - G_r(t)\right\} \\
&\quad + \sum_{r=1}^{k} p_r \sqrt{n_l} \int W(t) \left\{\hat{\Lambda}_n^{(l)}(t) - \Lambda_0(t)\right\} dG_r(t) \\
&\equiv \Delta_{1n}^{(l)} + \Delta_{2n}^{(l)} + \Delta_{3n}^{(l)} + \Delta_{4n}^{(l)}
\end{aligned}$$

for $l = 1, \ldots, k$. Using the same techniques as those used in the proof of Theorem 3.1 of Balakrishnan and Zhao (2009), we can show that $\Delta_{1n}^{(l)} = o_p(1)$, $\Delta_{2n}^{(l)} = o_p(1)$, and $\Delta_{3n}^{(l)} = o_p(1)$. Note that

$$\Delta_{4n}^{(l)} = \sqrt{n_l} \int A_l(t) \left\{\hat{\Lambda}_n^{(l)}(t) - \Lambda_0(t)\right\} dG_l(t) + o_p(1).$$

Then it follows from Theorem 1 that we have

$$\Delta_n^{(l)} = \bar{V}_n^{(l)} + o_p(1).$$

This proves Theorem 2.

Proof of Theorem 3. To prove $\hat{\sigma}_l^2 - \sigma_l^2 \to_p 0$, we first show that

$$\sup_{i\in S_l} \sup_{1\leqslant j\leqslant K_i^{(l)}} \left| \frac{\Delta G_n^{(r)}\left(T_{i,j}^{(l)}\right)}{\Delta G_n^{(l)}\left(T_{i,j}^{(l)}\right)} - \frac{g_r\left(T_{i,j}^{(l)}\right)}{g_l\left(T_{i,j}^{(l)}\right)} \right| \to_p 0, \quad r, l = 1, \ldots, k, \tag{A1}$$

where $\Delta G_n^{(l)}(t) = G_n^{(l)}(t) - G_n^{(l)}(t-)$. Set

$$A = \left\{ \min_{1\leqslant l\leqslant k} \min_{i\in S_l} \min_{1\leqslant j\leqslant K_i^{(l)}} \left(T_{i,j}^{(l)} - T_{i,j-1}^{(l)}\right) \geqslant s_0 \right\}.$$

From Condition A4, we have $\Pr(A)=1$. Choose $0<\varepsilon_0<1/2$ and $\delta=n^{-\varepsilon_0/2}$. Thus, for $n>s_0^{-2/\varepsilon_0}$ and $\omega\in A$, we have

$$\Delta G_n^{(r)}\big(T_{i,j}^{(l)}\big)=G_n^{(r)}\big(T_{i,j}^{(l)}\big)-G_n^{(r)}\big(T_{i,j}^{(l)}-\delta\big),$$
$$j=1,\ldots,K_i^{(l)},\quad i\in S_l,\quad l=1,\ldots,k.$$

Note that for $r=1,\ldots,k$ and $t\in[\tau_0,\tau]$,

$$\begin{aligned}G_n^{(r)}(t)-G_n^{(r)}(t-\delta)&=\big\{G_n^{(r)}(t)-G_r(t)\big\}+\{G_r(t)-G_r(t-\delta)\}\\&\quad+\big\{G_r(t-\delta)-G_n^{(r)}(t-\delta)\big\}\\&=o_p\big(n^{-\frac{1}{2}+\frac{\varepsilon_0}{2}}\big)+g_r(t)\delta+g_r'(\xi)\delta^2,\end{aligned}$$

where $\xi\in(t-\delta,t)$. Then, for $r,l=1,\ldots,k$, we have

$$\frac{\Delta G_n^{(r)}\big(T_{i,j}^{(l)}\big)}{\delta}=o_p\big(n^{-\frac{1}{2}+\varepsilon_0}\big)+g_r\big(T_{i,j}^{(l)}\big)+g_r'(\xi)\delta,$$
$$\xi\in\big(T_{i,j}^{(l)}-\delta,T_{i,j}^{(l)}\big)$$

and

$$\frac{\Delta G_n^{(r)}\big(T_{i,j}^{(l)}\big)}{\delta}-g_r\big(T_{i,j}^{(l)}\big)=o_p(n^{-c}),\tag{A2}$$

where L_0 is a positive constant and $0<c<\min(1/2-\varepsilon_0,\varepsilon_0/2)$. It follows from (A2) that

$$\begin{aligned}&\sup_{i\in S_l}\sup_{1\leqslant j\leqslant K_i^{(l)}}\left|\frac{\Delta G_n^{(r)}\big(T_{i,j}^{(l)}\big)}{\Delta G_n^{(l)}\big(T_{i,j}^{(l)}\big)}-\frac{g_r\big(T_{i,j}^{(l)}\big)}{g_l\big(T_{i,j}^{(l)}\big)}\right|\\&=\sup_{i\in S_l}\sup_{1\leqslant j\leqslant K_i^{(l)}}\left|\frac{\Delta G_n^{(r)}\big(T_{i,j}^{(l)}\big)/\delta}{\Delta G_n^{(l)}\big(T_{i,j}^{(l)}\big)/\delta}-\frac{g_r\big(T_{i,j}^{(l)}\big)}{g_l\big(T_{i,j}^{(l)}\big)}\right|\\&=o_p(n^{-c}).\end{aligned}$$

Next we will show that $\hat\sigma_l^2-\sigma_l^2=o_p(1)$ for $l=1,\ldots,k$. Define

$$\phi(\eta,\Lambda,\mathbf{X}^{(l)})=\sum_{j=1}^{K^{(l)}}\eta\big(T_j^{(l)}\big)\big\{N\big(T_j^{(l)}\big)-\Lambda\big(T_j^{(l)}\big)\big\}.$$

Then $\sigma_l^2=P_l\phi^2(A_l,\Lambda_0,\mathbf{X}^{(l)})$ and $\hat\sigma_l^2=P_n^{(l)}\phi^2(A_n^{(l)},\hat\Lambda_n^{(l)},\mathbf{X}^{(l)})$, where P_l is the probability measure of $\mathbf{X}^{(l)}$, $P_lf=\int f\,dP_l$, $P_n^{(l)}$ is the empirical measure corresponding to $\mathbf{X}^{(l)}$, and $P_n^{(l)}f=\sum_{i\in S_l}f(\mathbf{X}_i^{(l)})/n_l$. Note that

$$\begin{aligned}\hat\sigma_l^2-\sigma_l^2&=P_n^{(l)}\big\{\phi^2\big(A_n^{(l)},\hat\Lambda_n^{(l)},\mathbf{X}^{(l)}\big)-\phi^2\big(A_n^{(l)},\Lambda_0,\mathbf{X}^{(l)}\big)\big\}\\&\quad+P_n^{(l)}\big\{\phi^2\big(A_n^{(l)},\Lambda_0,\mathbf{X}^{(l)}\big)-\phi^2(A_l,\Lambda_0,\mathbf{X}^{(l)})\big\}\\&\quad+\big(P_n^{(l)}-P_l\big)\phi^2(A_l,\Lambda_0,\mathbf{X}^{(l)}).\end{aligned}$$

It can be easily shown that

$$P_n^{(l)}\big\{\phi^2\big(A_n^{(l)},\hat\Lambda_n^{(l)},\mathbf{X}^{(l)}\big)-\phi^2\big(A_n^{(l)},\Lambda_0,\mathbf{X}^{(l)}\big)\big\}=o_p(1)$$

and

$$\big(P_n^{(l)}-P_l\big)\phi^2(A_l,\Lambda_0,\mathbf{X}^{(l)})=o_p(1).$$

Note that

$$\begin{aligned}&\big|\phi\big(A_n^{(l)},\Lambda_0,\mathbf{X}_i^{(l)}\big)-\phi\big(A_l,\Lambda_0,\mathbf{X}_i^{(l)}\big)\big|=\big|\phi\big(A_n^{(l)}-A_l,\Lambda_0,\mathbf{X}_i^{(l)}\big)\big|\\&\leqslant\Big\{N_i^{(l)}\Big(T_{i,K_i^{(l)}}^{(l)}\Big)+\Lambda_0(\tau)\Big\}\sum_{j=1}^{K_i^{(l)}}\big|A_n^{(l)}\big(T_{i,j}^{(l)}\big)-A_l\big(T_{i,j}^{(l)}\big)\big|\\&\leqslant a_1\Big\{N_i^{(l)}\Big(T_{i,K_i^{(l)}}^{(l)}\Big)+\Lambda_0(\tau)\Big\}\\&\quad\times\sum_{j=1}^{K_i^{(l)}}\Bigg\{\big|W_n\big(T_{i,j}^{(l)}\big)-W(T_{i,j}^{(l)})\big|\\&\qquad+\max_{1\leqslant r\leqslant k}\left|\frac{\Delta G_n^{(r)}\big(T_{i,j}^{(l)}\big)}{\Delta G_n^{(l)}\big(T_{i,j}^{(l)}\big)}-\frac{g_r\big(T_{i,j}^{(l)}\big)}{g_l\big(T_{i,j}^{(l)}\big)}\right|\Bigg\}\end{aligned}$$

with probability 1 for some constant a_1 and

$$\begin{aligned}&\big|\phi\big(A_n^{(l)},\Lambda_0,\mathbf{X}_i^{(l)}\big)+\phi\big(A_l,\Lambda_0,\mathbf{X}_i^{(l)}\big)\big|=\big|\phi\big(A_n^{(l)}+A_l,\Lambda_0,X\big)\big|\\&\leqslant a_2K_i^{(l)}\Big\{N_i^{(l)}\Big(T_{i,K_i^{(l)}}^{(l)}\Big)+\Lambda_0(\tau)\Big\}\end{aligned}$$

with probability 1 for some constant a_2. These along with the Cauchy–Schwarz inequality, Conditions 1–3 and (5) and (6) give

$$\begin{aligned}&E\big|\phi^2\big(A_n^{(l)},\Lambda_0,\mathbf{X}_i^{(l)}\big)-\phi^2(A_l,\Lambda_0,\mathbf{X}_i^{(l)})\big|\\&\leqslant a_3E\Bigg[\Big\{N_i^{(l)}\Big(T_{i,K_i^{(l)}}^{(l)}\Big)+\Lambda_0(\tau)\Big\}^2\\&\qquad\times\Bigg\{\sum_{j=1}^{K_i^{(l)}}\big|W_n\big(T_{i,j}^{(l)}\big)-W\big(T_{i,j}^{(l)}\big)\big|\Bigg\}\Bigg]\\&\quad+a_3E\Bigg[\Big\{N_i^{(l)}\Big(T_{i,K_i^{(l)}}^{(l)}\Big)+\Lambda_0(\tau)\Big\}^2\\&\qquad\times\max_{1\leqslant r\leqslant k}\Bigg\{\sum_{j=1}^{K_i^{(l)}}\left|\frac{\Delta G_n^{(r)}\big(T_{i,j}^{(l)}\big)}{\Delta G_n^{(l)}\big(T_{i,j}^{(l)}\big)}-\frac{g_r\big(T_{i,j}^{(l)}\big)}{g_l\big(T_{i,j}^{(l)}\big)}\right|\Bigg\}\Bigg]\\&\leqslant a_3\Big[E\Big\{N_i^{(l)}\Big(T_{i,K_i^{(l)}}^{(l)}\Big)+\Lambda_0(\tau)\Big\}^4\Big]^{1/2}\\&\qquad\times\Bigg[E\Bigg\{\sum_{j=1}^{K_i^{(l)}}\big|W_n\big(T_{i,j}^{(l)}\big)-W\big(T_{i,j}^{(l)}\big)\big|\Bigg\}^2\Bigg]^{1/2}\\&\quad+a_3\Big[E\Big\{N_i^{(l)}\Big(T_{i,K_i^{(l)}}^{(l)}\Big)+\Lambda_0(\tau)\Big\}^4\Big]^{1/2}\\&\qquad\times\max_{1\leqslant r\leqslant k}\Bigg[E\Bigg\{\sum_{j=1}^{K_i^{(l)}}\left|\frac{\Delta G_n^{(r)}\big(T_{i,j}^{(l)}\big)}{\Delta G_n^{(l)}\big(T_{i,j}^{(l)}\big)}-\frac{g_r\big(T_{i,j}^{(l)}\big)}{g_l\big(T_{i,j}^{(l)}\big)}\right|\Bigg\}^2\Bigg]^{1/2}\\&\quad\longrightarrow 0,\end{aligned}$$

where a_3 is a finite positive constant. This completes the proof.